JOURNAL OF COMBINATORIAL THEORY, Series A 61, 173–192 (1992)

Qualitative Independence and Sperner Problems for Directed Graphs

L. GARGANO,* J. KÖRNER,† AND U. VACCARO*

Dipartimento di Informatica ed Applicazioni, Università di Salerno, 84081 Baronissi (SA), Italy

Communicated by the Managing Editors

Received June 30, 1990

Sperner's theorem about the largest family of incomparable subsets of an n-set is in fact a theorem about the largest anti-chain in the natural extension to sequences of a linear order. We replace the linear order by an arbitrary directed graph and ask for the cardinality of the largest set of incomparable sequences of length n one can form of the vertices. Two sequences are comparable if for every coordinate, all the arcs between corresponding vertices (if any) go in the same direction. Similarly, we look for the largest cardinality of sets of sequences that are incomparable in any graph from a given family. We find the asymptotic solution in some cases and give constructions in others. Our results imply new lower bounds on the cardinality of the largest family of qualitatively two-independent partitions in the sense of Rényi.

1. INTRODUCTION

Let us consider an arbitrary n-element set Y. Two partitions, P and P' of Y are called *qualitatively independent* if the number of non-empty classes in the partition they generate is equal to the product of the number of non-empty classes in the two partitions. More formally, if P has the non-empty classes $P_1, P_2, ..., P_k$ with $\bigcup_{i=1}^{k} P_i = Y$ while P' has the non-empty classes $P'_1, P'_2, ..., P'_{k'}$ with $\bigcup_{i=1}^{k'} P'_i = Y$, then P and P' are qualitatively independent if

$$P_i \cap P'_j \neq \emptyset \qquad \text{for every} \quad i \leqslant k, \quad j \leqslant k'.$$

This definition has an intuitive meaning in probability theory. Two partitions can be generated by two independent random variables if and only if the partitions are qualitatively independent.

* Work supported in part by the Italian Ministry of the University and of the Scientific Research in the framework of the "Algoritmi e Sistemi di Calcolo" project.

† On leave from the Mathematics Institute of HAS, 1364 Budapest, POB 127, Hungary.

0097-3165/92 $5.00

Let us consider only k-partitions, i.e., partitions into k classes. Let $N(n, k)$ be the largest cardinality of a family of k-partitions of an n-set under the restriction that any two partitions in the family are qualitatively independent. Poljak and Tuza [9] have shown that

$$\frac{\log k}{k(k-1)} \leqslant \limsup_{n \to \infty} \frac{1}{n} \log N(n, k) \leqslant \frac{2}{k}, \tag{1}$$

where the lower bound holds only if k is a prime power. The proof of the upper bound uses Bollobás' inequality [3], while the lower bound is obtained via projective geometries. No better upper bound is known, but Poljak and Tuza themselves have derived a better lower bound for $k = 3$. Writing

$$q_k = \limsup_{n \to \infty} \frac{1}{n} \log N(n, k),$$

their improved bound [9] is

$$q_3 \geqslant \tfrac{1}{3}.$$

In Körner and Simonyi [7] this is improved to

$$q_3 \geqslant 0.409....$$

In the present paper we strengthen the lower bound in (1) for small values of k. In particular, we further improve on the lower bound for q_3, showing by a novel construction that

$$q_3 \geqslant 0.483.... \tag{2}$$

We will obtain these results by the use of a construction technique applicable to a broader class of problems at the crossroads of combinatorics and information theory.

Sperner's beautiful 1928 theorem [12] about incomparable subsets of an n-set has a natural generalization in which the sets are replaced by n-length sequences of elements of a linearly ordered set [4]. The given order generates a partial order on the sequences in the obvious way and the maximum cardinality of an anti-chain in this partial order can be easily determined in terms of the rank function. To our knowledge, the analogous problem for partially ordered sets has never been posed. In fact, while the question remains the same, the answer to this more general problem seems to be quite different. In particular, there is no natural candidate for the largest set of what we call incomparable sequences. (Let us state again that we call two sequences incomparable if there are two coordinates in which their corresponding elements are ordered differently. Hence two sequences in which all the corresponding coordinates are unrelated in the partial order are not incomparable in our sense.) Furthermore, the problem does not become any easier if we only ask for asymptotic solutions, i.e., for the order of exponential growth of the largest set of incomparable sequences as their length tends to infinity.

Literally the same question can be asked for an arbitrary directed graph replacing the initial partial order. To illustrate the kind of problems we have in mind let us begin with a simple

EXAMPLE. Let T be the ternary alphabet $T=\{0, 1, 2\}$. Let G be the complete graph on T with the cyclic orientation in which the arcs are directed from 0 to 1, from 1 to 2, and from 2 back to 0.

The graph G defines a directed graph G_n on T^n in the natural way. There is an arc going from $\mathbf{x}$ to $\mathbf{x}'$ in G_n if for every i, either $x_i = x_i'$ or there is an arc in G going from x_i to x_i'; moreover $\mathbf{x} \neq \mathbf{x}'$. Two sequences, $\mathbf{x}$ and $\mathbf{x}'$ are called incomparable if they are not connected by a directed edge. Let $I(G, n)$ denote the maximum cardinality of a set of pairwise incomparable elements of G_n. (In other words, $I(G, n)$ is the stability number of G_n.) What is the asymptotics of $I(G, n)$ as n tends to infinity?

We have no real idea to tackle this question. We have chosen it to start because it is perhaps the best illustration of the conceptual simplicity and the mathematical difficulty of our subject.

Clearly, if $A \subset \{0, 1\}^n$ is the set of the characteristic vectors of a Sperner family of subsets of $\{1, 2, ..., n\}$, then A is an independent set in G_n. Hence

$$I(G, n) \geqslant \binom{n}{\lfloor n/2 \rfloor}.$$

On the other hand, trivially,

$$I(G, n) \leqslant 3^n.$$

No asymptotically significant improvement of these trivial bounds is known. More precisely, let us call

$$\Sigma(G) = \limsup_{n \to \infty} \frac{1}{n} \log I(G, n)$$

the *Sperner capacity* of G. The best bounds we know are

$$1 \leqslant \Sigma(G) \leqslant \log 3.$$

Note that here and in the sequel the logarithms are to the base 2.

The setup presented in the example has an immediate generalization for arbitrary directed graphs.

DEFINITION 1. Let G be an arbitrary directed graph with vertex set X. The sequences $\mathbf{x} \in X^n$ and $\mathbf{x}' \in X^n$ are *incomparable* for G, if

$$\exists i \ (x_i, x_i') \in E(G) \qquad \text{and} \qquad \exists j \ (x_j', x_j) \in E(G).$$

Let us call the set $A \subset X^n$ *incomparable* for G if the different sequences in A are incomparable.

Let $I(G, n)$ be the largest cardinality of an incomparable set in X^n. We call

$$\Sigma(G) = \limsup_{n \to \infty} \frac{1}{n} \log I(G, n)$$

the *Sperner capacity* of G.

Obviously, this definition is consistent with that in the example. The first basic question it raises is the relation of the Sperner capacity to the Shannon capacity of graphs.

DEFINITION 2 (Shannon [11]). Let G be an (undirected) graph with vertex set X. The graph G^n is defined on the vertex set X^n by the following edge set $E(G) \subset X^n$:

$$(\mathbf{x}, \mathbf{x}') \in E(G^n) \qquad \text{if} \quad \exists i \ (x_i, x_i') \in E(G).$$

Let $K(G, n)$ denote the largest cardinality of a complete subgraph in G^n. The quantity

$$C(G) = \limsup_{n \to \infty} \frac{1}{n} \log K(G, n)$$

is the Shannon capacity of G.

It is well known that in the above definition the limit superior can be replaced by a limit. More on the Shannon capacity of a graph can be found in Lovász [8].

For a directed graph G we can consider the corresponding undirected graph $\hat{G}$ in which $(x, x') \in E(\hat{G})$ if either (x, x') or (x', x) is a directed edge in G. Clearly,

$$\Sigma(G) \leqslant C(\hat{G}). \tag{3}$$

PROBLEM 1. Give an example for a graph G in which (3) does not hold with equality.

The upper bound in the example is a special case of (3). As the example shows, $\Sigma(G)$ is hard to find even for graphs the Shannon capacity of which is easily determined.

Although finding the Sperner capacity of a directed graph is a new problem in combinatorics, it is strongly related to an old one, that of finding the largest family of qualitatively independent partitions of an n-set. More precisely, a conceptually easy generalization of our previous problem contains that of qualitatively indepenent partitions as a special case. Next we state this generalization.

DEFINITION 3. Let $\mathscr{G}$ be a family of directed graphs each of which has vertex set X. Let us call the set $A \subset X^n$ *incomparable* for $\mathscr{G}$ if the different sequences in A are incomparable for every graph in $\mathscr{G}$. Let $I(\mathscr{G}, n)$ be the largest cardinality of an incomparable set in X^n. We call

$$\Sigma(\mathscr{G}) = \limsup_{n \to \infty} \frac{1}{n} \log I(\mathscr{G}, n)$$

the *Sperner capacity* of $\mathscr{G}$.

In this paper we will have little to say about Sperner capacity in the general case. Rather, our focus will be on qualitative independence. We will show, however, that the same technique can be applied to tackle many of the more general problems to which we shall return in more detail elsewhere.

The paper is in three parts. In the first part we describe our new construction of three-partitions yielding (2). In the second part we present a more general construction of the same kind that will allow us to improve on the lower bound in (1) for $k \leqslant 13$. In the last part of the paper we will then show that constructions of this kind give tight bounds for Sperner capacity in some special cases. This is noteworthy since the problem of qualitative independence is equivalent to Sperner capacity in some (other) special case.

2. QUALITATIVELY INDEPENDENT THREE-PARTITIONS

We show by an explicit construction that

THEOREM 1. $q_3 \geqslant 0.483\ldots$

Proof. For every n, we construct a "large" set of qualitatively independent three-partitions. Let us choose the n-element set $\{1, 2, \ldots, n\}$. We will represent any three-partition of this set by an element of $T^n = \{0, 1, 2\}^n$ in the obvious manner. Although there are six different ways of doing this, any of them suits us. In fact, the three-partitions represented by the ternary n-sequences $\mathbf{x}$ and $\mathbf{x}'$ are qualitatively independent if and only if for any $a \in \{0, 1, 2\}$, $b \in \{0, 1, 2\}$ there is a coordinate $i = i(a, b) = i(a, b, \mathbf{x}, \mathbf{x}')$ with $x_i = a$, $x_i' = b$, Hence, instead of constructing the partitions, we can construct a set of ternary sequences every pair of which satisfies the above nine conditions. In particular, it will follow that no partition is represented twice and thus the number of sequences equals the number of partitions we construct. Furthermore, we note that the three-conditions in which $a = b$ can be satisfied in a trivial manner, setting, e.g., $x_1 = 0$, $x_2 = 1$, $x_3 = 2$ in every sequence. Thus any set of sequences satisfying the remaining six conditions (involving $a \neq b$) can be modified to satisfy the original conditions by refixing the sequence 012 to each of its elements. Therefore, in order to prove the theorem, it is sufficient to construct a sequence of sets B_n satisfying all the above conditions for $a \neq b$ and such that

$$\limsup_{n \to \infty} \frac{1}{n} \log |B_n| \geqslant 0.483....$$

For brevity, let us call any set satisfying these six conditions a "good set." Similarly, if two sequences satisfy the conditions, we call them a "good pair."

Let us consider the set $A \subset T^*$ defined by

$$A = \{012, 120, 111212, 000220\}$$

and write $A_n = A^* \cap T^n$. In other words, A_n is the set of n-length ternary sequences one can form by concatenating in an arbitrary manner several repetitions of the four strings in A. By a well-known result of Shannon (cf. [6, Lemma 1.4.5]),

$$\limsup_{n \to \infty} \frac{1}{n} \log |A_n|$$

equals the unique positive root of the equation

$$2 \cdot 2^{-3\alpha} + 2 \cdot 2^{-6\alpha} = 1.$$

Hence

$$\limsup_{n \to \infty} \frac{1}{n} \log |A_n| \geqslant 0.483.... \tag{4}$$

We will show that A_n can be partitioned into polynomially (in n) many classes so that every class of the partition is a "good set" in our sense. In order to define the partition, we first introduce eight functionals on T^n. To this end, we need some new notation.

We shall refer to the elements of A as "words." We note first of all that a sequence $\mathbf{x} \in T^n$ can have at most one representation in the form of an element of A^*; i.e., it can be formally equal to at most one sequence of words from A. (This is immediate, since A is a prefix code; i.e., no word in A is a prefix of another one.) To define our functionals, let us define the set $I(\mathbf{a} \mid \mathbf{x})$ as the set of those coordinates i for which in the unique representation of $\mathbf{x}$ as an element of A^* the coordinate i is a starting position of a copy of the word $\mathbf{a}$. We set

$$N(\mathbf{a} \mid \mathbf{x}) = |I(\mathbf{a} \mid \mathbf{x})| \qquad \text{and} \qquad L(\mathbf{a} \mid \mathbf{x}) = \sum_{i \in I(\mathbf{a} \mid \mathbf{x})} i \tag{5}$$

for every $\mathbf{a} \in A$. This defines eight functionals on T^n. We partition T^n according to their different values on the sequences. More precisely, two sequences belong to the same class of our partition if all the above eight functionals take the same value on both of them.

We claim that every class of the partition is a "good set." This will prove the theorem, because at least one of the classes is sufficiently large. In fact, notice that for the integer-valued functionals $N(\cdot \mid \cdot)$ and $L(\cdot \mid \cdot)$ we have

$$0 \leqslant N(\mathbf{a} \mid \mathbf{x}) \leqslant n, \qquad 0 \leqslant L(\mathbf{a} \mid \mathbf{x}) \leqslant \frac{n^2}{2}, \qquad \text{for every} \quad \mathbf{a} \in A, \quad \mathbf{x} \in T^n,$$

and therefore, by a very rough estimate, our partition has at most $n^8(n+1)^4$ many classes. Let B_n be a class of the partition having largest cardinality. Then,

$$|B_n| \geqslant (n+1)^{-12} \, |A_n|.$$

Comparing this with (4) we see that

$$\limsup_{n \to \infty} \frac{1}{n} \log |B_n| \geqslant 0.483....$$

Hence it remains to prove that every class of our partition is a "good set," indeed. For this purpose, let us write

$$\mathbf{b} = 012$$

$$\mathbf{c} = 120$$

$$\mathbf{d} = 111$$

$$\mathbf{e} = 212$$

$$\mathbf{f} = 000$$

$$\mathbf{g} = 220.$$

The set A can be written as $A = \{\mathbf{b}, \mathbf{c}, \mathbf{de}, \mathbf{fg}\}$. We claim that any two different elements $\mathbf{x}$ and $\mathbf{x}'$ of A_n, on which the eight functionals (5) take the same values, form a good pair; i.e., for every $a \neq b$ from $\{0, 1, 2\}$ there is a coordinate i with $x_i = a$, $x_i' = b$.

We distinguish four cases:

Case 1. Suppose that in the two sequences $\mathbf{x}$ and $\mathbf{x}'$ all the six-letter words are in the same coordinates, i.e.,

$$I(\mathbf{a} \mid \mathbf{x}) = I(\mathbf{a} \mid \mathbf{x}') \qquad \text{if} \quad \mathbf{a} = \mathbf{de} \quad \text{or} \quad \mathbf{a} = \mathbf{fg}. \tag{6}$$

Since

$$N(\mathbf{b} \mid \mathbf{x}) = N(\mathbf{b} \mid \mathbf{x}') \qquad N(\mathbf{c} \mid \mathbf{x}) = N(\mathbf{c} \mid \mathbf{x}'),$$

and by our hypothesis, every occurrence of $\mathbf{b}$ in $\mathbf{x}$ is matched by an occurrence of either $\mathbf{b}$ or $\mathbf{c}$ in the same positions in $\mathbf{x}'$, it follows that there must be an occurrence of $\mathbf{b}$ in $\mathbf{x}$ matched by an occurrence of $\mathbf{c}$ in $\mathbf{x}'$ in the same coordinates, and vice versa, there also must be an occurrence of $\mathbf{b}$ in $\mathbf{x}'$ matched by a $\mathbf{c}$ in $\mathbf{x}$ in the same coordinates. Noting that all the pairs of elements of $\{0, 1, 2\}$ occur in matched positions in the two sequences $\mathbf{b}$ and $\mathbf{c}$, we conclude that $\mathbf{x}$ and $\mathbf{x}'$ form a good pair.

There are two cases in which (6) holds for one of the six-letter words from A but not for the other. These cases can be dealt with in essentially the same way.

Case 2. Let us suppose first that

$$I(\mathbf{fg} \mid \mathbf{x}) = I(\mathbf{fg} \mid \mathbf{x}')$$

but the same is not true for $\mathbf{de}$. On the other hand, $\mathbf{de}$ occurs in $\mathbf{x}$ and $\mathbf{x}'$ the same number of times. Let us denote by $I_m(\mathbf{a} \mid \mathbf{x})$ the beginning coordinate of the mth occurrence of $\mathbf{a}$ in the sequence $\mathbf{x}$ (read as a sequence from A^*.) (If $m > N(\mathbf{a} \mid \mathbf{x})$, we put $I_m(\mathbf{a} \mid \mathbf{x}) = 0$.) By our assumptions, there is some m such that

$$I_m(\mathbf{de} \mid \mathbf{x}) \neq I_m(\mathbf{de} \mid \mathbf{x}').$$

Hence there must be some m for which

$$I_m(\mathbf{de} \mid \mathbf{x}) < I_m(\mathbf{de} \mid \mathbf{x}') \tag{7}$$

or else we would not have $L(\mathbf{de} \mid \mathbf{x}) = L(\mathbf{de} \mid \mathbf{x}')$. Moreover, if we consider the smallest integer m for which (7) holds, then, necesarily, we also have

$$I_m(\mathbf{de} \mid \mathbf{x}) \in I(\mathbf{b} \mid \mathbf{x}') \cup I(\mathbf{c} \mid \mathbf{x}');$$

i.e., the **d** of the mth occurrence of **de** in **x** is matched with **b** or **c** in **x**$'$. Namely, in the contrary case it could only be matched with some part of a **de** in **x**$'$. Moreover, this then would be the lth **de** in **x**$'$ for some $l<m$, implying

$$I_l(\mathbf{de}\mid\mathbf{x})<I_l(\mathbf{de}\mid\mathbf{x}'),$$

a contradiction. In conclusion, we find that

$$x_j=1, \quad x_j'=0, \qquad x_k=1, \quad x_k'=2, \tag{8}$$

for some j and k among the three consecutive coordinates beginning with $I_m(\mathbf{de}\mid\mathbf{x})$. Reversing the role of **x** and **x**$'$ in the above, we find coordinates with

$$x_s=0, \quad x_s'=1, \qquad x_t=2, \quad x_t'=1. \tag{9}$$

Looking at the sequences from right to left rather than from left to right as before, we shall find the missing coordinates. In fact, let m now be the *largest* integer for which (7) holds. Like before, we must have

$$I_m(\mathbf{de}\mid\mathbf{x}')+3\in I(\mathbf{b}\mid\mathbf{x})\cup I(\mathbf{c}\mid\mathbf{x}),$$

as by our assumption, the mth is the first-from-the-right unmatched occurrence of **de** in **x**$'$ for which it precedes (fom the right) the corresponding occurrence of **de** in **x**. Hence the **e**-part of **de** can only be matched with **b** or **c** in **x**. We see that the simultaneous occurrence of an **e** in **x**$'$ and either a **b** or a **c** in **x** guarantees a coordinate such that

$$x_u=0, \qquad x_u'=2. \tag{10}$$

Reversing the role of **x** and **x**$'$ in the preceding argument gives a coordinate v with

$$x_v'=2, \qquad x_v=0.$$

This, (8), (9), and (10) establish our claim.

Case 3. Suppose that

$$I(\mathbf{de}\mid\mathbf{x})=I(\mathbf{de}\mid\mathbf{x}')$$

but the same is not true for **fg**. The proof is analogous to that in the preceding case. In fact, we now see that a simultaneous occurrence of **f** with either **b** or **c** guarantees the 0–1 and 0–2 coordinate pairs, whereas a **g** matched with either of **b** or **c** guarantees the 2–1 coordinate pair. With these modifications the proof for the previous case literally applies.

Case 4. We have seen that $\mathbf{x}$ and $\mathbf{x}'$ form a good pair unless

$$I(\mathbf{de} \mid \mathbf{x}) \neq I(\mathbf{de} \mid \mathbf{x}'), \qquad I(\mathbf{fg} \mid \mathbf{x}) \neq I(\mathbf{fg} \mid \mathbf{x}'). \tag{11}$$

This is therefore the only remaining case to examine. We proceed indirectly. Suppose first that

$$x_i = 1, \qquad x_i' = 0 \tag{12}$$

never occurs. This means that if $j \in I(\mathbf{de} \mid \mathbf{x})$, then the sequence $x_j' x_{j+1}' x_{j+2}'$ equals either $\mathbf{d}$ or $\mathbf{e}$. It follows that

$$I_m(\mathbf{de} \mid \mathbf{x}) \geqslant I_m(\mathbf{de} \mid \mathbf{x}') \qquad \text{for every } m,$$

a contradiction. Hence we obtain the existence of a coordinate i that gives (12). Next suppose that

$$x_k = 2, \qquad x_k' = 1 \tag{13}$$

never occurs. A similar argument as before shows that if $x_j x_{j+1} x_{j+2}$ is the sequence $\mathbf{g}$; i.e., the coordinate $j-3$ is in $I(\mathbf{fg} \mid \mathbf{x})$; then the sequence $x_j' x_{j+1}' x_{j+2}'$ equals either $\mathbf{f}$ or $\mathbf{g}$, and thus

$$I_m(\mathbf{fg} \mid \mathbf{x}) \leqslant I_m(\mathbf{fg} \mid \mathbf{x}'),$$

once again contradicting (11). Therefore, we obtain a coordinate k with property (13).

Finally, suppose that

$$x_l = 2, \qquad x_l' = 0 \tag{14}$$

is not to be found. This implies that every $\mathbf{e}$ in $\mathbf{x}$ is matched by either a $\mathbf{d}$ or an $\mathbf{e}$ in $\mathbf{x}'$, yielding

$$I_m(\mathbf{de} \mid \mathbf{x}) \leqslant I_m(\mathbf{de} \mid \mathbf{x}') \qquad \text{for every } m,$$

the usual contradiction with (11). In conclusion, we have found coordinates i, k, and l such that (12)–(14) hold. Reversing the role of $\mathbf{x}$ and $\mathbf{x}'$ makes the proof complete. ∎

3. More on Qualitative Independence

The construction technique applied in the proof of Theorem 1 has its roots in [7]. Similar constructions often yield non-trivial bounds for problems we can interpret in terms of Sperner capacities. The core of the technique is a lemma that we now state in its general form that is suitable for different applications.

For an arbitrary finite set X, let $l(\mathbf{x})$ denote the length of a sequence $\mathbf{x} \in X^*$. Given aritrary sequences $\mathbf{x}, \mathbf{x}' \in X^*$, we denote by $B(\mathbf{x}, \mathbf{x}')$ the set consisting of those ordered pairs $(a, b) \in X^2$, for which $(x_i, x_i') = (a, b)$ holds in at least one coordinate i of $\mathbf{x} = x_1 \cdots x_{l(\mathbf{x})}$ and $\mathbf{x}' = x_1 \cdots x_{l(\mathbf{x}')}$. If $l(\mathbf{x}) \neq l(\mathbf{x}')$, then different positions of the two sequences will coincide according to whether the two sequences have a coinciding initial coordinate or a coinciding last coordinate, respectively (as parts of some larger sequences located in matched-at-the-beginning, resp. matched-at-the-end positions.) Therefore, we define $E(\mathbf{x}, \mathbf{x}')$ as the set of those ordered pairs $(a, b) \in X^2$ for which $(x_{l(\mathbf{x})-i}, x'_{l(\mathbf{x}')-i}) = (a, b)$ in at least one coordinate, i.e., for some $0 \leqslant i \leqslant \min\{l(\mathbf{x}), l(\mathbf{x}')\}$.

LEMMA 1 (Two-words lemma). *Given two sequences,* $\mathbf{a} \in X^*$, $\mathbf{b} \in X^*$ *such that none is a prefix of the other, there exists a sequence of sets* $B_n \subset A^* \cap X^n$*—where* $A = \{\mathbf{a}, \mathbf{b}\}$*—with the following properties*:

$$\limsup_{n \to \infty} \frac{1}{n} \log |B_n| = \limsup_{n \to \infty} \frac{1}{n} \log |A^* \cap X^n|, \tag{15}$$

$$B(\mathbf{a}, \mathbf{b}) \subseteq B(\mathbf{x}, \mathbf{x}'), \qquad E(\mathbf{a}, \mathbf{b}) \subseteq B(\mathbf{x}, \mathbf{x}') \qquad \textit{for every} \quad \mathbf{x} \in B_n, \quad \mathbf{x}' \in B_n \quad \textit{with} \quad \mathbf{x} \neq \mathbf{x}'. \tag{16}$$

Remark. Obviously, $B(\mathbf{x}, \mathbf{x}') = E(\mathbf{x}, \mathbf{x}')$. The lemma holds for every uniquely decipherable code with two words A, i.e., every pair $(\mathbf{a}, \mathbf{b})$ with the property that any sequence X^* has at most one decomposition into a sequence of elements of A. We have chosen the more restrictive formulation to keep the statement of the lemma simpler. Nevertheless, the proof below literally applies under the more general condition as the property of prefix codes we use is just their unique decipherability.

Proof. As in the proof of the previous theorem, we observe that every sequence $\mathbf{x} \in X^*$ can be written as an element of A^* in at most one way. Thus we can define the numbers $N(\mathbf{c} \mid \mathbf{x})$ and $L(\mathbf{c} \mid \mathbf{x})$ for $\mathbf{c} = \mathbf{a}$, $\mathbf{c} = \mathbf{b}$ as in the previous proof. We partition $A_n = A^* \cap X^n$ once again according to the different values of the quadruple

$$\{N(\mathbf{a} \mid \mathbf{x}), N(\mathbf{b} \mid \mathbf{x}), L(\mathbf{a} \mid \mathbf{x}), L(\mathbf{b} \mid \mathbf{x})\}.$$

Let B_n be a class of this partition with largest cardinality. As in the previous proof, we have

$$\limsup_{n \to \infty} \frac{1}{n} \log |B_n| = \limsup_{n \to \infty} \frac{1}{n} \log |A_n|.$$

Hence it suffices to prove that B_n has properties (15)–(16). To this end, let us define $I_m(\mathbf{a} \mid \mathbf{x})$ as the first coordinate of the mth occurrence of $\mathbf{a} \in A$ in the unique representation of $\mathbf{x}$ in the form of an element of A^*. Consider any $\mathbf{x}, \mathbf{x}' \in B_n$ with $\mathbf{x}' \neq \mathbf{x}$. Since the two sequences are different, but $L(\mathbf{a} \mid \mathbf{x}) = L(\mathbf{a} \mid \mathbf{x}')$, there must be some m_0 such that

$$I_{m_0}(\mathbf{a} \mid \mathbf{x}) < I_{m_0}(\mathbf{a} \mid \mathbf{x}'). \tag{17}$$

Let m denote the smallest integer m_0 for which (17) holds. If the two sequences coincide to the left of the coordinate $I_m(\mathbf{a} \mid \mathbf{x})$, we conclude that the A^*-representation of $\mathbf{x}'$ has a $\mathbf{b}$ in the position starting at the coordinate $I_m(\mathbf{a} \mid \mathbf{x})$, and thus

$$B(\mathbf{a}, \mathbf{b}) \subseteq B(\mathbf{x}, \mathbf{x}').$$

Next we show that this inclusion always holds. In fact, we will show that $\mathbf{x}'$ has a $\mathbf{b}$ starting in the position $I_m(\mathbf{a} \mid \mathbf{x})$ even if $\mathbf{x}$ and $\mathbf{x}'$ disagree somewhere to the left of said coordinate. Let us look at the coordinate $I_m(\mathbf{a} \mid \mathbf{x})$ of $\mathbf{x}'$. Suppose, indirectly, that in the unique A^*-representation of $\mathbf{x}'$ this coordinate belongs to the lth occurrence of $\mathbf{a}$. In this hypothesis, we cannot have $l \geqslant m$, or else m would not satisfy (17). But we cannot have $l < m$, either, for the latter would imply that already $m-1$ satisfies (17), in contrast with our hypothesis that m is the smallest number with this property. We conclude that the coordinate $I_m(\mathbf{a} \mid \mathbf{x})$ of $\mathbf{x}'$ belongs to an occurrence of $\mathbf{b}$. Let $\mathbf{x}'_-$ denote the part of the sequence $\mathbf{x}'$ that precedes this $\mathbf{b}$, and let $\mathbf{x}'_+$ denote the sequence $\mathbf{x}'$ up to the last coordinate of the same $\mathbf{b}$. (Thus the sequence $\mathbf{x}'_+$ ends with this $\mathbf{b}$). Further, let $\mathbf{x}_-$ denote the beginning segment of length $I_m(\mathbf{a} \mid \mathbf{x}) - 1$ of $\mathbf{x}$. We have to prove $l(\mathbf{x}'_-) = l(\mathbf{x}_-)$. We have

$$N(\mathbf{a} \mid \mathbf{x}'_-) = N(\mathbf{a} \mid \mathbf{x}_-)$$

by the definition of m. Hence the indirect hypothesis $l(\mathbf{x}'_-) < l(\mathbf{x}_-)$ implies that

$$N(\mathbf{b} \mid \mathbf{x}'_-) < N(\mathbf{b} \mid \mathbf{x}_-).$$

This means that

$$N(\mathbf{b} \mid \mathbf{x}'_+) \leqslant N(\mathbf{b} \mid \mathbf{x}_-)$$

while we still have

$$N(\mathbf{a} \mid \mathbf{x}'_+) = N(\mathbf{a} \mid \mathbf{x}_-).$$

Therefore, $l(\mathbf{x}'_+) \leqslant l(\mathbf{x}_-) < I_m(\mathbf{a} \mid \mathbf{x})$, which contradicts

$$l(\mathbf{x}'_+) \geqslant I_m(\mathbf{a} \mid \mathbf{x}).$$

We conclude that the coordinate $I_m(\mathbf{a} \mid \mathbf{x})$ is a simultaneous starting point for an **a** in **x** and a **b** in **x**′. Hence

$$B(\mathbf{a} \mid \mathbf{b}) \subseteq B(\mathbf{x}, \mathbf{x}').$$

An analogous argument applied from the right end of the sequences implies that there is another coordinate which, in turn, is a simultaneous endpoint for an **a** in **x** and a **b** in **x**′. Thus

$$E(\mathbf{a}, \mathbf{b}) \subseteq B(\mathbf{x}, \mathbf{x}'). \quad \blacksquare$$

The Two-Words lemma (TWL) allows us to improve on the Poljak–Tuza lower bound (1) for $k \leqslant 13$. As this method is clearly sub-optimal, we will not try to apply it in the best-possible way. Rather, we will limit ourselves to showing that there is an easy way to improve on the bound (1). As is illustrated by Theorem 1, similar but more complicated constructions based on sets of more than two words in the role of A might yield substantial further improvements.

THEOREM 2. *Let $t(a, b)$ denote the unique non-negative root of the equation*

$$2^{-at} + 2^{-bt} = 1 \tag{18}$$

for arbitrary integers $a > 0$, $b > 0$. We have

$$q_4 \geqslant t(3, 5), \qquad q_5 \geqslant t(5, 6), \qquad q_6 \geqslant t(8, 10)$$

$$q_7 \geqslant t(11, 12), \qquad q_8 \geqslant t(14, 18), \qquad q_9 \geqslant t(18, 19)$$

$$q_{10} \geqslant t(25, 26), \qquad q_{11} \geqslant t(28, 29), \qquad q_{12} \geqslant t(36, 37)$$

$$q_{13} \geqslant t(39, 40).$$

Proof. We start with $k = 4$. Let us apply the TWL for the following set $A = \{\mathbf{a}, \mathbf{b}\}$:

$$\mathbf{a} = 012$$

$$\mathbf{b} = 12330.$$

We see that

$$B(\mathbf{a}, \mathbf{b}) = \{(0, 1), (1, 2), (2, 3)\}$$

$$E(\mathbf{a}, \mathbf{b}) = \{(0, 3), (1, 3), (2, 0)\}.$$

Hence the union

$$B(\mathbf{a}, \mathbf{b}) \cup E(\mathbf{a}, \mathbf{b}) \cup B(\mathbf{b}, \mathbf{a}) \cup E(\mathbf{b}, \mathbf{a})$$

contains all the 12 pairs of unequal elements from $\{0, 1, 2, 3\}$. In virtue of the TWL there exists a sequence of sets $B_n \subset A^* \cap X^n$ with

$$\limsup_{n \to \infty} \frac{1}{n} \log |B_n| = \limsup_{n \to \infty} \frac{1}{n} \log |A^* \cap X^n| \tag{19}$$

with the property that every unequal pair $\mathbf{x}, \mathbf{x}'$ of elements of B_n also contains the above 12 pairs from $\{0, 1, 2, 3\}$. Applying Shannon's capacity theorem [6, Lemma 1.4.5] we see by (19) that

$$\limsup_{n \to \infty} \frac{1}{n} \log |B_n| = t(l(\mathbf{a}), l(\mathbf{b})) = t(3, 5).$$

Representing four-partitions of $\{1, 2, 3, ..., n\}$ by elements of $\{0, 1, 2, 3\}^n$ analogously to the way we proceed in the proof of Theorem 1 we conclude that

$$q_4 \geqslant t(3, 5).$$

The remaining proofs are exactly the same. The key point is to find, for every k in question, two words, $\mathbf{a} \in \{0, 1, ..., k-1\}^*$ and $\mathbf{b} \in \{0, 1, ..., k-1\}^*$ such that $B(\mathbf{a},\mathbf{b}) \cup E(\mathbf{a}, \mathbf{b})$ contain all the pairs of unequal elements of $\{0, 1, ..., k-1\}$ in at least one of their two orderings. The rest follows automatically.

For $k = 5$ choose

$$\mathbf{a} = 0\ 1\ 2\ 3\ 4$$

$$\mathbf{b} = 1\ 2\ 3\ 4\ 0\ 1$$

We have $l(\mathbf{a}) = 5$, $l(\mathbf{b}) = 6$.

If $k = 6$, choose

$$\mathbf{a} = 0\ 1\ 2\ 3\ 4\ 5\ 0\ 1$$

$$\mathbf{b} = 1\,2\,3\,4\,5\,0\,2\,3\,4\,5.$$

The word-lengths are 8 and 10.

If $k = 7$, the words

$$\mathbf{a} = 0\ 1\ 2\ 3\ 4\ 5\ 6\ 5\ 6\ 0\ 1$$

$$\mathbf{b} = 1\ 2\ 3\ 4\ 5\ 6\ 0\ 1\ 2\ 3\ 4\ 5$$

have the lengths 11 and 12.

For $k=8$, we choose

$$\mathbf{a}=0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 7\ 0\ 1\ 2\ 5\ 6$$

$$\mathbf{b}=1\ 2\ 3\ 4\ 5\ 6\ 7\ 0\ 1\ 2\ 3\ 4\ 3\ 4\ 5\ 6\ 7\ 0,$$

with lengths 14 and 18, respectively.

In the case $k=9$, the choices are

$$\mathbf{a}=0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 7\ 8\ 0\ 1\ 2\ 3\ 4\ 5\ 6$$

$$\mathbf{b}=1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 0\ 1$$

with the lengths 18 and 19.

In order to describe the remaining constructions, we introduce, for $r<s$, the shorthand notation $[r, s]$ for the sequence $r, r+1, ..., s$. With this notation, we can describe our last constructions as follows:

For $k=10$ our choice is

$$\mathbf{a}=[0,9]\ [8,9]\ [0,7]\ [6,9]\ 0$$

$$\mathbf{b}=[1,9]\ [0,9]\ [0,6]$$

of lengths 25 and 26, respectively.

If $k=11$, we write

$$\mathbf{a}=[0,10]\ [9,10]\ [0,8]\ [7,10]\ [0,1]$$

$$\mathbf{b}=[1,10]\ [0,10]\ [0,7]$$

with respective lengths 28 and 29.

We define for $k=12$,

$$\mathbf{a}=[0,11]\ [10,11]\ [0,11]\ [0,6]$$

$$\mathbf{b}=[1,11]\ [0,11]\ 0\ [0,11]\ 0$$

with respective lengths 36 and 37.

Finally, we have for $k=13$ the choices

$$\mathbf{a}=[0,12]\ [11,12]\ [0,12]\ [0,10]$$

$$\mathbf{b}=[1,12]\ [0,12]\ 0\ [3,12]\ [0,3]$$

with $l(\mathbf{a})=39$, $l(\mathbf{b})=40$. ▌

A comparison with the corresponding values in the Poljak–Tuza bound (1) shows that in all the above cases we obtain some improvement. An analysis of why this happens and why it stops happening for larger values of k will be one of the subjects of the next section.

4. Sperner Capacities

Let us return to the Sperner capacity problem of Definition 3. It is clear from the foregoing that q_k is a Sperner capacity in the sense of Definition 3. The corresponding family $\mathscr{G}$ of graphs is that of the one-edge graphs defined by the different edges of the complete graph on k vertices.

A particularly intriguing special case of the problem of Definition 3 arises, in fact, precisely when all the individual graphs in the family $\mathscr{G}$ have a single edge. It is clear that in this particular case the orientation of the edges has no importance. Let us denote by $\mathscr{F}(G)$ the family of all the one-edge graphs the different members of which correspond to the different edges of a given (non-directed) graph G. We reformulate this special case of the Sperner capacity problem for the reader's convenience. We recall that two sequences, $\mathbf{x}, \mathbf{x}' \in [V(G)]^n$ are incomparable for $\mathscr{F}(G)$ if for every $(a, b) \in E(G)$ there exist coordinates i and j such that

$$(x_i, x_i') = (x_j, x_j') = (a, b),$$

where, e.g., x_i stands for the ith coordinate of $\mathbf{x}$. Let $J(G, n)$ be the largest cardinality of any set $C \subset [V(G)]^n$, every two distinct elements of which are incomparable for $\mathscr{F}(G)$. We write

$$\Theta(G) = \limsup_{n \to \infty} \frac{1}{n} \log J(G, n).$$

In our previous language, for every edge of G, the family $\mathscr{F}(G)$ has a graph with vertex set $V(G)$ and an edge set consisting of this single edge. Clearly,

$$\Theta(G) = \Sigma(\mathscr{F}(G)), \tag{20}$$

where $\Sigma(\mathscr{F}(G))$ is the Sperner capacity of the family of graphs $\mathscr{F}(G)$); let us recall that formally, the family must consist of directed graphs, but in case of single edge-graphs any orientation would lead to the same notion, and therefore, we will not specify any.

We will show that in two special cases the technique presented in the previous section solves the capacity problem by exhibiting optimal or asymptotically optimal constructions. One of these cases does not involve any new result. In fact, it has been proved by Körner and Simonyi [7] that

Theorem KS. *For the graph G with*

$$V(G) = \{0, 1, 2\}, \qquad E(G) = \{(0, 1), (0, 2)\},$$

we have

$$\Theta(G) = \log \frac{1+\sqrt{5}}{2}.$$

We will not repeat the proof here. Let us just outline the construction establishing

$$\Theta(G) \geqslant \log \frac{1+\sqrt{5}}{2},$$

since it is immediate from our TWL. (This is no surprise. We have already mentioned that the TWL is a generalization of the proof technique used in [7]). To see that the last inequality holds, apply the TWL to

$$\mathbf{a} = 0$$

$$\mathbf{b} = 12$$

and observe that

$$B(\mathbf{a}, \mathbf{b}) \cup E(\mathbf{a}, \mathbf{b}) = \{(0, 1), (0, 2)\}.$$

Let us turn to

LEMMA 2. *Let C_k be the cycle on k vertices. We have*

$$\Theta(C_k) \geqslant t\left(\left\lceil \frac{k}{2} \right\rceil, \left\lceil \frac{k}{2} \right\rceil + 1\right).$$

Proof. We can write

$$V(C_k) = \{0, 1, ..., k-1\}$$

$$E(C_k) = \{(i, i+1); i = 0, 1, ..., k-1\},$$

where the numbers are understood modulo k. We will apply the TWL with the choice

$$\mathbf{a} = 1, 3, 5, ..., 2\left\lfloor \frac{k-1}{2} \right\rfloor + 1$$

$$\mathbf{b} = 0, 2, 4, ..., 2\left\lfloor \frac{k-1}{2} \right\rfloor, 0.$$

One easily sees that

$$B(\mathbf{a},\mathbf{b}) \cup E(\mathbf{a}, \mathbf{b}) \cup B(\mathbf{b}, \mathbf{a}) \cup E(\mathbf{b}, \mathbf{a}) = \bigcup \{(a, b) \cup (b, a)\},$$

where the union is extended to all the edges (a, b) of C_k. ∎

The following upper bound is an easy consequence of the results of [5]. Its proof is based on elementary information theory.

LEMMA 3. $\Theta(C_k) \leqslant 2/k$.

Proof. One easily sees that

$$\Theta(C_k) \leqslant \max_P \min_{(a,b) \in E(C_k)} [P(a)+P(b)]\, h\left(\frac{P(a)}{P(a)+P(b)}\right),$$

where in the maximization P is running over all the probability distributions on the vertices of C_k, and $h(x)$ is the binary entropy function:

$$h(x) = -x \log x - (1-x)\log(1-x).$$

Clearly, the maximum in the above inequality is achieved if P is the uniform distribution on the vertices of C_k. ∎

If the reader prefers not to fill in the missing simple details, a more detailed version of this argument is available in [5].

The two preceding lemmas immediately give the following

THEOREM 3. $\lim_{k \to \infty} (k\Theta(C_k)/2) = 1$.

Proof. We have

$$t\left(\left\lceil \frac{k}{2} \right\rceil, \left\lceil \frac{k}{2} \right\rceil + 1\right) \leqslant \Theta(C_k) \leqslant \frac{2}{k}.$$

Obviously,

$$\frac{2}{k+2} \leqslant t\left(\left\lceil \frac{k}{2} \right\rceil, \left\lceil \frac{k}{2} \right\rceil + 1\right).$$

Hence

$$\frac{2}{k+2} \leqslant \Theta(C_k) \leqslant \frac{2}{k}. \quad ∎$$

We have seen that the TWL gives reasonably sharp estimates for the $\Theta(C_k)$. A more elaborate construction of the same kind should yield similar results for other families of graphs.

We conclude the paper with a discussion of how $\Theta(C_k)$ relates to q_k. The union of two graphs, $F \cup G$ is understood in the sense

$$V(F \cup G) = V(F) \cup V(G)$$

$$E(F \cup G) = E(F) \cup E(G).$$

A primitive way of guaranteeing the simultaneous presence of various edges of a given graph between sequences is to assign different subsets of coordinates to different subgraphs the union of which is the given graph. More precisely, we have

LEMMA 4. *If $G = G_1 \cup G_2$, then*

$$\Theta(G) \geqslant \frac{\Theta(G_1)\Theta(G_2)}{\Theta(G_1) + \Theta(G_2)}.$$

Proof. One easily sees that for every $\alpha \in (0, 1)$,

$$\Theta(G) \geqslant \max_{\alpha} \ \min\{\alpha\Theta(G_1), (1-\alpha)\Theta(G_2)\}. \tag{21}$$

In fact, consider a sequence of constructions $A_{\alpha n} \subset [V(G_1)]^{\alpha n}$ achieving $\Theta(G_1)$ and a sequence of constructions $B_{(1-\alpha)n} \subset [V(G_2)]^{(1-\alpha)n}$ achieving $\Theta(G_2)$. Suppose, without loss of generality that

$$\alpha\Theta(G_1) > (1-\alpha)\Theta(G_2).$$

Consider a arbitrary subset of $A_{\alpha n}$ of cardinality $|B_{(1-\alpha)n}|$. Upon establishing a one to one correspondence between the respective elements of $B_{(1-\alpha)n}$ and those of the chosen subset of $A_{\alpha n}$, we obtain a construction of length n, cardinality $|B_{(1-\alpha)n}|$, and such that every pair of its elements satisfies all the conditions imposed by the graph $G_1 \cup G_2$.
The construction is obtained upon juxtaposition of the corresponding shorter sequences satisfying the constraints for G_1 and G_2, respectively. This proves (21). The lemma follows by choosing

$$\alpha = \frac{\Theta(G_2)}{\Theta(G_1) + \Theta(G_2)}. \quad \blacksquare$$

Note that by a well-known theorem of graph theory (cf. [1, Corollary 1, p. 230]), for k odd, the complete graph K_k can be decomposed into $(k-1)/2$ edge-disjoint copies of C_k, the cycle of length k. Applying the previous lemma to this decomposition, we obtain

COROLLARY 1. *For every odd k,*

$$q_k = \Theta(K_k) \geqslant \frac{2}{k-1}\,\Theta(C_k).$$

This yields by Theorem 2 the estimate

$$q_k \geqslant \frac{2}{k-1}\,\Theta(C_k) \geqslant \frac{4}{(k-1)(k+2)},$$

for odd k. For small k, this estimate already improves on the Poljak–Tuza lower bound (1). The gain is possible because for small k the roughness of the estimate in Corollary 1 is compensated by our rather precise evaluation of $\Theta(C_k)$. The constructions in Theorem 2 are based on a somewhat better "glueing together" of those for individual cycles than the primitive juxtaposition suggested in Lemma 4.

Many questions remain. It is striking that the known upper bounds for $\Theta(K_k)$ and $\Theta(C_k)$ are the same. In view of Theorem 3, especially because of the last inequality in the proof, it is tempting to conjecture that $\Theta(C_k) = 2/k$ for every k.

ACKNOWLEDGMENTS

We are grateful to Gábor Simonyi for sharing his ideas with us. Thanks are due to an anonymous referee for pointing out the right result on cycle decompositions of the complete graphs and other helpful remarks.

Note added in proof. We have proved the conjecture that $q_k = 2/k$ in the paper "Sperner Capacities, Graphs and Combinatorics," to appear. Problem 1 has been solved by Calderbank, Graham, Sheep, Frankl, and Li in the paper "The Sperner Capacity of the cyclic triangle for linear and nonlinear codes," manuscript, 1992.

REFERENCES

1. C. Berge, "Graphs," North-Holland, Amsterdam, 1985.
2. C. Berge, "Hypergraphes," Dunod, Paris, 1988.
3. B. Bollobás, On generalized graphs, *Acta Math. Acad. Sci. Hungar.* **16** (1965), 441–452.
4. N. G. de Bruijn, C. Tengbergen, and D. Kruyswijk, On the set of divisors of a number, *Nieuw Arch. Wisk.* **23** (1951), 191–193.
5. G. Cohen, J. Körner, and G. Simonyi, Zero-error capacities and very different sequences, *in* "Sequences: Combinatorics, Compression, Security and Transmission" (R. M. Capocelli, Ed.), pp. 144–155 Springer-Verlag, New York/Berlin, 1990.
6. I. Csiszár and K. Körner, "Information Theory. Coding Theorems for Discrete Memoryless Systems," Academic Press, New York, 1982.
7. J. Körner and G. Simonyi, A Sperner-type theorem and qualitative independence, *J. Combin. Theory Ser. A* **59** (1992), 90–103.
8. L. Lovász, On the Shannon capacity of a graph, *IEEE Trans. Inform. Theory* **IT-25**, No. 1 (1979), 1–7.
9. S. Poljak and Z. Tuza, On the maximum number of qualitatively independent partitions, *J. Combin. Theory Ser. A* **51** (1989), 111–116.
10. A. Rényi, "Foundations of Probability," Wiley, New York, 1971.
11. C. E. Shannon, The zero-error capacity of a noisy channel, *IRE Trans. Inform. Theory* **2** (1956) 8–19.
12. E. Sperner, Ein Satz über Untermengen einer endlichen Menge, *Math. Z.* **27** (1928), 544–548.